AD-A094 837 SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX F/G 6/16
FORTY-EIGHT VERSUS TWENTY-FOUR HOUR DUTY FOR USAF MISSILE CREWS--ETC(U)
NOV 80 S F GRAY
UNCLASSIFIED SAM-TR-80-39 NL

1 OF 1
AD A
094837

END
DATE
FILMED
3-81
DTIC

AD A094837

Report SAM-TR-80-39

(12)

LEVEL II

FORTY-EIGHT VERSUS TWENTY-FOUR HOUR DUTY FOR USAF MISSILE CREWS: A FEASIBILITY STUDY USING SUBJECTIVE MEASURES OF FATIGUE

Stephen F. Gray, Major, USAF

DTIC
ELECTE
FEB 10 1981
E

November 1980

Final Report for Period August 1978 - March 1980

Approved for public release; distribution unlimited.

USAF SCHOOL OF AEROSPACE MEDICINE
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas 78235

DDC FILE COPY

81 2 10 025

NOTICES

This final report was submitted by personnel of the Crew Performance Branch, Crew Technology Division, USAF School of Aerospace Medicine, Aerospace Medical Division, AFSC, Brooks Air Force Base, Texas, under job order 7930-10-28.

When U.S. Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The operational personnel who participated in this study were fully briefed on all procedures prior to participation in the study.

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

STEPHEN F. GRAY, Major, USAF
Project Scientist

WILLIAM F. STORM, Ph.D.
Supervisor

ROY L. DEHART
Colonel, USAF, MC
Commander

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

14

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: SAM-TR-80-39

2. GOVT ACCESSION NO.: AD-A094837

3. RECIPIENT'S CATALOG NUMBER: 9

4. TITLE (and Subtitle): 6 FORTY-EIGHT VERSUS TWENTY-FOUR HOUR DUTY FOR USAF MISSILE CREWS: A FEASIBILITY STUDY USING SUBJECTIVE MEASURES OF FATIGUE.

5. TYPE OF REPORT & PERIOD COVERED: Final Report, Aug[redacted]78 - Mar[redacted]80

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): 10 Stephen F. Gray, Major, USAF

8. CONTRACT OR GRANT NUMBER(s): 17 1∅

9. PERFORMING ORGANIZATION NAME AND ADDRESS:
USAF School of Aerospace Medicine (VNE)
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas 78235

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62202F 16 7930/10-28

11. CONTROLLING OFFICE NAME AND ADDRESS:
USAF School of Aerospace Medicine (VNE)
Aerospace Medical Division (AFSC)
Brooks Air Force Base, Texas 78235

12. REPORT DATE: 11 Nov[redacted]80

13. NUMBER OF PAGES: 54

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 12 158

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Minuteman
Missile Crews
Subjective Fatigue
Work/Rest Schedules
Extended Work Periods

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The purpose of this study was to assess the feasibility of implementing a 48-hour work schedule for missile launch crews of a United States Air Force operational missile wing.

A 90-day field test using two operational Minuteman missile squadrons as test and control groups was accomplished during the winter of 1978-1979 at Grand Forks Air Force Base, North Dakota.

DD FORM 1 JAN 73 1473

317000

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT (continued)

Subjective reports of fatigue and the quantity and quality of sleep were recorded daily by crew members during work and while they were off duty. In addition, subjective reports of workload and disruptions of sleep were gathered during alerts.

Subsets of data from at most 10 distributed alerts for each crew member were subjected to analyses of variance. The primary analyses were tests for differences between the responses of crew members working a new 48-hour vs. those working the standard 24-hour schedule. Scaled subjective reports of fatigue at the end of alert, after driving back to base, and after 24 hours of recovery indicated no significant differences between the two work schedules. Tests for interactions of the work schedules with activities during alerts, specific control centers, and cumulative or maturational effects over several alerts showed no effects that might have masked real differences between the work schedules.

A statistically significant difference between work schedules was found in the hours slept at home during the first night of recovery. Members of the test group (48-hour work schedule) increased their sleep during recovery by a mean value of 2.1 hours over the quantity of sleep on the night previous to the alert. Members of the control group (24-hour work schedule) increased their sleep during recovery by a mean value of 1.3 hours over the pre-alert quantity, a significantly smaller change than that of the test group.

This finding was interpreted as the result of having a greater "opportunity" to sleep which the test group was provided, rather than from a greater "need" to sleep resulting from the strain associated with the schedule. Support for this position was drawn from the failure to find significant differences between the groups in other variables measured.

Several threats to validity that existed in this field test were discussed. These included nonrandom assignment of treatment to groups, differential treatment of test and control groups in certain aspects of the procedures, the possibility that data were biased by the failure of crew members to complete all response materials, and other possible failures of control procedures that may have confounded the results. The success or failure of methods implemented to deal with these threats to validity was discussed.

In conclusion, a recommendation was made that the 48-hour work schedule was feasible under the specific circumstances used in this study.

TABLE OF CONTENTS

Accession For
NTIS GRA&I X
DTIC TAB
A

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (Cont)

FORTY-EIGHT VERSUS TWENTY-FOUR HOUR DUTY FOR USAF MISSILE CREWS: A FEASIBILITY STUDY USING SUBJECTIVE MEASURES OF FATIGUE

INTRODUCTION

Personnel assigned to military nuclear weapons systems work in unique environments and man their stations around-the-clock on schedules that are considerably different from most civilian jobs.

The United States Air Force operates a constantly "ready" deterrent force of intercontinental ballistic missiles (ICBMs) armed with nuclear payloads. These weapons systems are manned by launch crews working in underground control centers at remote locations in the United States. The Air Force must employ these crews efficiently to assure that the "readiness" mission is accomplished using limited human and material resources. To more efficiently use these resources, a new work schedule for manning missile control centers was proposed.

The 90-day field evaluation was conducted to evaluate the effect of the proposed work schedule on the crew members' health and safety. The 321st Strategic Missile Wing, Grand Forks Air Force Base, North Dakota, requested the USAF School of Aerospace Medicine, the Air Force laboratory responsible for biomedical research and development, to assist their unit in evaluating the effect of the new schedule on Minuteman ICBM launch crew members.

Because readers may not be familiar with military systems, especially intercontinental ballistic missiles, aspects of the missile crews' job and environment that are important to the evaluation of their work schedule are described. Following this description the literature relevant to the possible methods of evaluating the effect of the work schedule is reviewed, and the design of the present study is described.

Missile System

The U.S. Air Force's Titan and Minuteman ICBMs, deployed since the early 1960s, are maintained in around-the-clock readiness in underground silos safe from nuclear attack. This description of missile operations is concentrated on the Minuteman ICBM system, the subject of this study, but discussion of the Titan missile system is included where similarities and differences are informative. Two-member Minuteman launch crews operate small, blast-proof, underground control centers. In their work environment the crews are surrounded by equipment which is used to monitor, operate, and control 10 remotely located missiles. Additional equipment and devices in their environment provide redundant communications, electrical power, and environmental control. The control center is equipped with a food storage area, a bunk for one man, and toilet facilities. This environment is characterized by close quarters, constant noise, constant light, vibration, and cool dry air. The primary activity of the crew is the visual and auditory monitoring of equipment

PRECEDING PAGE BLANK-NOT FILMED

for information regarding the status of the missiles. Much of their time is spent communicating over radios and land lines to personnel in the above ground support facility, maintenance crews, security teams, other control center crews, and the staff at the base command post.

Many tasks require the participation of both crew members, but at other times one may sleep while the other remains awake. Prior to 1977, nuclear safety rules required both crew members to remain awake to monitor each other. Because of this requirement, a second crew was housed in the aboveground facility and the crews worked alternate 12-hour shifts. In 1977 modifications of launch control equipment eliminated the need for monitoring each other's activities. This change allowed one man to sleep and eliminated the need for a second crew above ground (39). Since that time the standard duty has been 24 hours. This period of duty, called an "alert tour" or an "alert," runs from approximately noon of the first day, when the crew assumes responsibility for the control center from the off-going crew, until noon of the following day when they are relieved by a fresh crew. An important aspect of Minuteman operations is the fact that the control centers are located some distance from the main base. The crews drive long distances (over 200 miles in many cases) in military vehicles from the main base to their assigned control centers. The elapsed time from the predeparture briefing at the beginning of the alert day to the debriefing following their return usually ranges from 30 to 36 hours. The six Minuteman missile wings are located at Air Force bases in the north central and northwestern states where winter weather is often severe, making travel slow and difficult.

In addition to the 7 or 8 alerts that standard crews work each month, missile units require their crews to perform training, office administration, and other duties in accordance with a master schedule for the entire unit. Some of the nonalert duties, such as training in the missile procedures trainer, must be done in the evening and on weekends.

The U.S. Air Force treats the control of nuclear weapons as a very serious and important responsibility. Missile crews' tasks and procedures while on alert are standardized by preprinted checklists and written directives. Crew members are constantly trained and retrained in procedures, and evaluated on their knowledge and proficiency. Missile units are under the scrutiny of the headquarters' staff who conduct frequent exercises, inspections, and evaluations to assess the integrity of the missile forces. The Air Force Flight Surgeon's Guide (1) sums up the possible problems and effects of missile duty on missile crew members:

> In missile operations, the relative influence of stress and environmental factors upon the individual is somewhat different from that in the flying situation. Stresses that do exist, such as the hazards of explosion, noxious fumes, and accidents, are, in general, less acute and anxiety-producing. Chronic combat tensions, obviously, are unlikely. Environmental factors, on the other hand, are of much greater significance. Remote locations, small stations with few personnel, limited recreational facilities, insufficient and possibly substandard housing are all important variables.

> The strategy of deterrence depends upon the awesome destructive potential of our missiles. Should we have to launch them, then our strategy has failed. Ability to operate the weapon system with very short warning is essential for the success of this strategy. The constant vigilance which is necessary demands alert, keyed-up people.
>
> Maintaining such a frame of mind is difficult enough in a hot war where use of the weapon system is imminent. It is a prodigious task to overcome complacency and establish an alert attitude in the cold war situation where the weapon system has failed in its mission if it has to be operated. These factors--remote location, cold war, and lack of opportunity to operate the system--all tend to render an individual more susceptible to impaired efficiency from emotional symptoms.
>
> Furthermore, it has been shown that from 25% to 40% of all missile failures are caused by human error. A sizeable proportion of these errors is due to impaired efficiency from underlying emotional tensions. Momentary lapses of attention, simple mistakes, slipshod and careless work can often be directly traced to emotional pressures. (p. 9-5)

Standard Work Schedule

The 24-hour alert schedule has been the standard for the 4-member, Titan missile crews since that system was first deployed, and as described earlier, this schedule has been standard for Minuteman crews since 1977. The 24-hour schedule, with its mid-day changeover, assures that most travel is accomplished during daylight. Because round-trip travel time to some control centers is more than 8 hours, the 24-hour schedule provides balance between travel and working time. Shorter schedules would not provide either of these advantages, and any schedule that requires more than one daily trip would also require more personnel and more fuel. The number of Minuteman crews was reduced by more than one-third when the 24-hour alert was introduced in 1977 (39).

Missile crew personnel complained of having very little useable time off during the month when they were working eight 24-hour alerts in addition to the required duty days for training and administrative matters (40, 42). Crews that worked at the most distant control centers, particularly in Minuteman units where the control centers are extremely remote, said that travel time used up a great deal of their "return" day. Complaints of this nature along with an interest in reducing travel costs led to proposals for instituting 48-hour alerts that would cut travel time and costs in half. In addition, it would permit intervals of more consecutive days between the longer alerts than were possible on the 24-hour schedule.

Proposed Work Schedule

A test of the 48-hour schedule was accomplished in April and May of 1977 at a Titan missile unit (42). The test involved 8 crews working for 2 months at a launch control center that had been modified to reduce noise and vibration. Crew members kept diaries, completed open-ended questionnaires, reported on their sleep, and were observed by safety and quality control evaluators during their alerts. Conclusions reported at the end of this evaluation were favorable to the 48-hour schedule with the stipulation that noise suppression would be accomplished at all control centers. The report of this evaluation discussed individual differences in reponse to the schedule and stated that the method of study may have changed the alert situation enough to make generalizing to other situations difficult; 48-hour alerts have not been implemented in Titan missile operations.

Minuteman crews generally must travel greater distances to their control centers than Titan crews, and severe weather is more often found at Minuteman wings than at the three Titan wings in Arizona, Kansas, and Arkansas. Minuteman wings also became interested in reducing travel time and cost, and a pilot test similar to the one previously described was accomplished using volunteers at Grand Forks Air Force Base (40). The results of this test were more favorable than the Titan study and did not rely on noise suppression as a prerequisite for implementating the new schedule. Both the Titan and Minuteman studies examined sleep, feelings of fatigue, noise, vibration, and concerns about safety while driving to the base after the long alert. The Titan test also emphasized performance on the job. Standardization, safety, and maintenance evaluators made special evaluations of the Titan crews during the test alerts. Both Titan and Minuteman tests included questionnaires that required the participants to judge their ability to perform the wartime mission during the extended alerts.

Proposed Study

When the 321st Strategic Missile Wing requested support from the USAF School of Aerospace Medicine in evaluating a long-term, large-scale test of 48-hour alerts, they emphasized their concern for "crew effectiveness" and the ability of the crews to accomplish their mission (25, 38).

Hartman (22) discussed the need for a quantitative basis for decisions regarding operational effectiveness:

> The operational goal which we must address is to determine on a quantitative basis the impact of mission duration/mission cycling on mission effectiveness. The qualitative aspects of this problem are well known to operational users of weapon systems. In practice, users develop rules and schedules for employing crews in weapons systems based on this qualitative understanding of crew limitations, refine these through experience, and arrive at workable compromise between mission requirements and crew capabilities. The goal set by operational managers is to avoid crew fatigue and loss of efficiency. System

designers borrow from operational experience in designing new weapons systems. The results at both the design and operational levels are, however, best guesses. A quantified description of crew capability in the area of mission duration/mission cycling will improve the coupling of crew duty-time limits and weapons system capability based on operational criteria. (p. 13-3)

Hartman was referring primarily to air operations, but what he said also applies to the problems of scheduling missile operations.

Government managers rely on methods such as operations research, systems analysis, economic analysis, and quantitative analysis to assist them in decision making. These types of analyses for managing resources are in common use from the Office of Management and Budget, through the Department of Defense, to operational missile units. Planning and decision-making criteria for missile operations are often described quantitatively under headings such as "missile-in-commission rates," "crew proficiency failure rates," and "crew-to-missile ratios." In developing the field evaluation of the 48-hour schedule for Minuteman operations, an attempt was made to use quantitative judgment criteria similar to the processes commonly used in decision making by missile wing commanders and managers.

Correspondence from the 321st Strategic Missile Wing (25, 38, 41) defined their concerns differently each time the proposed assessment was discussed. At one point they specified that the crews' ability to execute the Emergency War Order mission was the primary concern. Crew effectiveness was stressed as the important issue at another point. In the final test plan (41) seven factors were specified for evaluation.

1. Crew performance
2. Crew member preference and comments
3. Physiological and psychological effects of the extended alert on crew members
4. Transportation costs
5. Ground safety
6. Effects of adverse weather on crew members and scheduling
7. Family attitudes

The test plan further specified that the USAF School of Aerospace Medicine would "study and evaluate the impact of the extended alert tour on crew fatigue, stress, and ground safety" (p. 2).

Most of the Wing's correspondence relating to the crews' performance, safety, and health emphasized the problem of fatigue. Therefore the USAF School of Aerospace Medicine considered systematically measuring fatigue and describing its effects on the crew members participating in this test, to be the primary method of assisting the 321st Strategic Missile Wing in their endeavor.

Literature Review

Reviews of research on stress and fatigue in military airlift operations (9, 22, 23) provided much of the background for the review of literature on fatigue which follows. Reviews of fatigue by Cameron (7) and Grandjean (18) provided continuity between military and civilian interests in the study of fatigue. Symposia, conferences, and meetings regarding stress and fatigue called by the military and by industry provided useful references and guidance to other literature. Experience as a Titan missile launch crew member and as a field investigator in a previous study of the change to 24-hour alerts in Minuteman operations (39) provided the present author with first-hand knowledge of the stresses of missile duty and the effects of those stresses on the missile crews. Literature concerned with noise, vibration, sleep deprivation, sleep disturbance, driving fatigue, and circadian variation in man was reviewed because these are present in the missile crew members' job.

Definitions and Measures--The term "fatigue" is commonly used to describe physical or mental weariness or exhaustion resulting from exertion. Authors writing about fatigue as a subject of scientific study usually have defined or described fatigue beyond its use in common language. Browne (6) reported that Mosso in 1884 had described muscular and mental fatigue and recommended that methods of measuring fatigue should be devised. Often the methods proposed by an investigator for measuring fatigue were emphasized when he defined or described it. McFarland (30) said that depending on the interests and background of those attempting to measure fatigue, it is usually described by leading adjectives such as physiological, psychological, clinical, operational, or performance-related. He stated, "There is probably no single word in our vocabulary which has been less adequately described or understood, yet few people would deny personal acquaintance with it" (p. 1).

Methods of measuring fatigue can be classified into three types: physiological, behavioral, and subjective. One or more of these three methods are accepted by most fatigue researchers as important to understanding fatigue, but the researchers disagree about which measures to use to best characterize fatigue. Ash (2) defined fatigue in these lengthy terms:

> Fatigue is a comprehensive term which in its widest application embraces all those immediate and temporary changes whether of a functional or organic character, which take place within an organism or any of its constituent parts as a direct result of its exertion, and which tend to interfere with or inhibit the organism's further activities. Its principal effect is loss of efficiency, a lessening of the capacity to do or sustain activity; its most obvious sign is depression--a lowering of sensitivity so that a given stimulus calls forth a response of less magnitude and intensity after exertion than before. (p. 1)

He attempted to demonstrate fatigue by showing decrements in muscular or mental control by subjects performing specific physical or mental tasks.

Tasks developed to measure fatigue became known as fatigue tests. Muscio (33) suggested that fatigue tests were of no use in the study of work, and instead researchers should study the work itself and its physiological effects. Further he said the the word "fatigue" should be abolished from scientific discussion because it was imprecise. Muscio (34) recommended instead that the "feeling-tone," an introspective report of workers' feelings of fitness, should be measured and compared to work production curves.

Sullivan (43) asked workers to describe their subjective mood by having them mark a point on a line representing a continuum from low to high cheerfulness. She also measured their pulse, blood pressure, and muscular strength. These measures were in agreement with Muscio's recommendations, but she also presented a battery of tests of mental function that included visual memory, tapping performance, color naming, adding numbers, and free association. She attempted to relate all measured variables to fatigue in the workers she studied.

Poffenberger (37) designed a seven-point scale for subjective report of fatigue and compared these reports to performance on four mental tests. While subjective reports all showed that fatigue increased from the initial level, performance on some of the mental tests improved, and on others it decreased. This study has often been cited to indicate that subjective feelings of fatigue are poor predictors of performance.

In discussing subjective reports in relation to behavioral measures, Bartlett (4) said, "Almost all the earlier scientific investigators--Mosso, Kraepelin, Rivers, Thorndike--have reported, with evidence, that almost any such subjective statement is consistent with almost any type of performance" (p. 1). He proposed instead a definition of fatigue from his research with simulated flying tasks in the 1940s.

> Fatigue is a term used to cover all those determinable changes in the expression of an activity which can be traced to the continuing exercise of that activity under its normal operational conditions, and which can be shown to lead either immediately or after delay, to deterioration in the expression of that activity, or, more simply, to results within the activity that are not wanted. (p. 1)

Bartlett's emphasis was similar to Ash's (2); both relied on the decrement in performance on a continuous task as a measure of fatigue. Bartlett reported that the subjective feelings of discomfort experienced by subjects in these tasks came too late (after performance had deteriorated) to be useful in studying fatigue.

Bartley and Chute (5) categorized variables from previous fatigue research into three groups: Physiological impairment or incapacity, "work decrement" from causes other than impairment, and subjective feelings of lassitude and disinclination towards activity. They emphasized the qualitative, individual feelings of fatigue and suggested studying those feelings by introspection. They stated that common sense and common experience of fatigue were important to its study. In their investigations of fatigue, Bartley and Chute were not concerned with decrements in performance because they said these decrements usually recover quickly with rest.

Some tasks such as vigilance or "watch keeping" require little physical or mental effort, yet they result in feelings of fatigue. Several authors (10, 14, 18, 44, 47, 49) discussed the importance of arousal in sedentary work. Colquhoun explained the idealized "inverted-U" shaped curve that Duffy's "activation" theory described. This is typically shown with performance graphed on the ordinate and activation on the abscissa. When activation is very low or not present, performance is poor because of errors of omission. As activation increases, performance improves up to an optimal point, beyond which performance degrades as a result of errors of commission. The activation in this idealized situation could be provided by the task, the environment, or the subject.

In a review of fatigue for an Ergonomics symposium on the subject, Grandjean (18) presented what he called "a simplified scheme of the neurophysiological concept of fatigue" (p. 428). Activation theory was used to explain some of the contradictory findings such as fatigue without effort and the reversal of performance decrement by novel stimuli. Grandjean concluded:

> In light of the present neurophysiological knowledge, we may consider fatigue as a state of the central nervous system controlled by the activity of the inhibitory and activating system of the brain stem. The regulating systems in turn are susceptible to reaction to stimuli from the surrounding world, to stimuli from the conscious part of the brain, and to humoral factors originating within the organism and having obviously the task of regulating recovery and wakefulness. The state of fatigue is accompanied by a decrease in *motivation to work*, a decrease in physical and mental performances, and by the occurrence of subjective feelings of fatigue. The latter induce animals and human beings to behavior-ensuring recovery. (pp. 435-436)

Investigators who utilized activation theory to explain fatigue and deficits in performance often relied on physiological measures of the activity of the central and autonomic nervous systems (14). Grandjean used critical flicker fusion and biochemical analysis of urine; others have employed the electrocardiogram, electroencephalogram, electromyogram, and galvanic skin response. Hartman et al. (23) and Hartman (22) described a model of the nervous system similar to Grandjean's and suggested that the interrelationships among measures of performance, subjective feelings, and neuroendocrine activity are to be expected because of their control by functionally related areas of the reticular formation, hypothalamus, and other subcortical structures.

Dukes-Dobos (15) rejected both performance decrement and subjective feelings as fatigue measures because they can exist even when no effort is expended. Instead he defined fatigue as:

> A normal psychophysiological process, which starts immediately after the beginning of any physical or mental activity and consists of the utilization of the body's energy stores, the accumulation of the breakdown products, and the activation of adaptive mechanisms which maintain the homeostasis of the organism. (p. 31)

Dukes-Dobos proposed measuring anabolic and catabolic products excreted in the urine as the only measure of fatigue. This approach contrasted with that of Hartman et al. (23), who recommended assessing fatigue using a complementary battery of measures that included performance, biochemistries, and subjective feelings.

Cameron (7) stated in discussing other author's definitions of fatigue:

> Fatigue is thus a concept which defies precise definition. It is a useful label for a generalized response to stress over a period of time, which has identifiable and measurable characteristics, but it has no explanatory value. It is not legitimate to describe any change in the individual's behavior as "due" to fatigue, since the term is no more than a general description of his personal state at the time such changes are noted. (p. 640)

Cameron reviewed the history of research on fatigue and pointed out three periods of interest in measuring fatigue. Early research was concerned with the relationship of fatigue and work output in industrial settings. This was followed in the 1940s by concern for safety of operations in aviation. Cameron's third historical period of interest focused on driving fatigue which has much in common with aviation fatigue in its concern with accidents.

Cameron described fatigue in industrial production and during task performance in aviation as short-term effects because rest is usually sufficient for recovery. In addition, he said that only when the effects of fatigue are cumulative over time are they interpretable and useful for predicting problems. Cameron recommended measuring arousal or activation by some physiological method at the time of stress and after a period of recovery. He theorized that the best method of quantifying fatigue was to measure recovery time. He offered this method as a way of comparing different types of working conditions.

Hartman (20) demonstrated that tne duration of sleep during recovery from fatigue is a useful index for studying the stresses of military operations. Although Hartman relied on subjective reports to measure this variable rather than the physiological methods proposed by Cameron, the basic purpose of each of the two methods is the same, i.e., to measure the time required to recover from an event or condition that has already ended.

Measuring fatigue in various environments and work situations has often been a difficult and disappointing enterprise. Researchers are usually interested in the effects on performance of a particular situation which is thought to be stressful or a cause of fatigue. Direct measures of performance are often difficult to interpret, because the measurements change the situation being measured. Similar problems are found using physiological measures that are believed to correlate with performance. The methods of measuring are often invasive to the subject or at least intrusive on the situation, so their results are also difficult to interpret. Most often performance and physiological measures are confined to laboratory studies where adequate experimental controls often remove many of the problems found in on-the-job settings. Unfortunately the results of much of this laboratory research cannot be

applied to actual work situations because the laboratory situations are distantly removed from the real world by the controls used.

Subjective Measures--Subjective measures have often been used in laboratory situations to provide a common basis among measures and because they can be measured in field studies to provide a common base of comparison between laboratory and field situations.

Weybrew (48), in a review of military research, noted that psychological and psychiatric measures were used most often in long-term field studies, whereas measures of performance and physiological effects were most often used in laboratory and simulation studies.

Advantages of Subjective Measures--The electroencephalogram has been used in sleep deprivation research (24, 28, 46) to assess ongoing effects and recovery from the deprivation period. Subjective reports have correlated well with these physiological measures. Thayer (44) discussed the importance of physiological measures of autonomic nervous system activity for assessing activation or arousal. He noted in using these physiological methods that there were great individual differences among subjects and temporal divergence among the measures. He proposed that a battery or composite of such measures was necessary to properly account for the effects of idiosyncratic and temporal variation. Thayer proposed that a method of controlled self report would provide comparable measures across subjects. He used the Activation-Deactivation Adjective Checklist and concurrently measured physiological indicators of activity of the autonomic nervous system and found four subjective factors that correlated closely with the physiological measures. He proposed using this easily measured, paper-and-pencil instrument in place of the complex equipment required to measure physiological variables.

Several researchers who have investigated the activation continuum using behavioral and physiological measures also have relied on subjectively reported feelings to verify their results. Eason and Dudley (16) reported that all these measures and the simultaneous environmental changes are needed to assess the intensity and direction of the activation continuum in a given situation. Thayer (44) stated, "phenomenological awareness of total bodily functioning . . . may be more representative of general bodily activation than any single peripheral physiological system" (p. 677). Subjects' self reports have been described as a summing up of the overall psychophysiological state. Mohler (31) compared feelings of fatigue experienced within the individual to "the subjective manner in which the sensation of thirst is experienced" (p. 238). Innes (26) called fatigue, reported by aircrew members, "a summary of certain aspects of their current situation. It is a diagnosis" (p. 5-1).

Advantages of using subjective measures in both laboratory and field settings are the ease of measurement, the adaptability to uses in various situations, and the fact they they provide a composite or even holistic approach to human measurement.

Subjective measures have disadvantages also. They are prone to motivational problems similar to those discussed in measuring performance. Another problem is that they are in fact under the control of the subject. Subjects are usually aware of how to respond in order to bring about a desired outcome

of the research. Often they respond in a particular way to provide the investigator the outcome they think he expects. Sometimes they choose not to respond at all.

One of the complaints against scaled subjective measures is that they depend on words. Words are ambiguous; they mean different things to different people. The same state of fatigue may be described differently by different persons. The factors of concern in measuring fatigue are defined by words. An example of the difficulties that arise follows: Pearson (35) designed a subjective instrument for measuring fatigue. In describing it, he said it "should, in toto, reflect the subjective state for any fatigue-research situation conceivable" (p. 191). Yet Innes described Pearson's checklist as a measure of "exhaustion" fatigue and attempted to develop a similar measure to isolate "nervous" fatigue. Innes used a layout and response methodology similar to Pearson's, but he used different words.

Wolf (49) attempted to develop measures of three factors or types of fatigue using a single instrument, a checklist of symptoms which subjects selected by a three-choice Likert response scale to describe their feelings. Wolf designed the checklist to investigate the interaction of motivation and sedentary work. He described three factors: "drowsy" fatigue, "nervous" fatigue, and "exhaustion" fatigue, and attempted to measure each of these by responses to items on the checklist that loaded on each of the three factors. Wolf used two sedentary pursuit-rotor tasks and a hand-tapping task, intending to elicit the three types of fatigue. Wolf found that short exposures to the tasks produced decrements in performance in approximately half his subjects, that dichotomous measures of effort did not predict responses to the fatigue factors, and that drowsy and exhaustion fatigue were not independent of each other. Both these factors were independent of nervous fatigue.

The Subjective Fatigue Checkcard--Pearson (35), using methods from attitude scaling, developed a short Feeling-Tone Checklist that represented a Guttman unidimensional scale; i.e., knowing the response to a specific item on the scale, one can predict the score. The method of responding to the scale is a three-choice Likert intensity method: Statements on the checklist are judged by the subject to be "worse than," "same as," or "better than" the subject's present feelings.

Pearson validated the checklist in laboratory tests using Air Force airmen as subjects. Five-hour sessions on a complex perceptual-motor task employing a multidimensional pursuit test apparatus were used to fatigue subjects in a test group. Fatigue scores from this test group were compared to a group of control subjects, who were kept alert during a similar period by being in an environment simulating that of jet pilots in "alert status" awaiting a call to man their planes. The scores easily discriminated the two groups although both groups' scores indicated increased fatigue over the test period. Pearson and Byars (36) examined the efficacy of using the Feeling-Tone Checklist to assess changes to affective states caused by drugs. An analeptic (Dexedrine), a depressant (Benadryl-hyoscine), or a placebo were provided randomly to subjects who were observed for 4 1/2 hours in a situation similar to the control group in the previously described experiment. Differences in responses to the checklist among the three treatments were significant and in the expected directions for the drugs used. Pearson

proposed using the Feeling-Tone Checklist to assess fatigue in industrial settings, operational field research, and studies testing the effects of drugs.

The Feeling-Tone Checklist has been reduced from 13 to 10 items for use by the USAF School of Aerospace Medicine as the Subjective Fatigue Checkcard. It has been used in numerous laboratory and field studies in which fatigue was of interest.

Although the attempt by Wolf (49) to develop a three-factor measure of fatigue was not successful, his findings indicate that the unidimensional continuum of the Subjective Fatigue Checkcard may be sufficient for measuring fatigue that results from both uneventful sedentary work and physical work. The Subjective Fatigue Checkcard does not provide responses that describe symptoms of nervousness, tension, or irritability that were discussed by several authors (5, 12, 18, 24, 26, 49). These symptoms of fatigue are expected to result from mental tasks that are associated with high activation. Although the items on the checkcard do not load on a factor of nervous fatigue, the checkcard has measured variations in fatigue resulting from high activation mental tasks. Three studies using the Subjective Fatigue Checkcard are reviewed here to indicate the diversity of its use and the support it gains from concurrent measures.

Harris et al. (19) reported that scores from the Subjective Fatigue Checkcard correlated with oral temperature during baseline recording, inflight activities, and recovery in an evaluation of aircrews flying long-term cargo operations. These researchers noted that the subjective measures followed the circadian rhythm of oral temperature during the normal baseline, then shifted with the physiological measure when the temperature rhythm was disrupted during rapid changes of time-zones on flights across the Pacific Ocean.

Cushman (13) reported using Pearson's original Feeling-Tone Checklist (13 items) to measure general fatigue in an experiment that compared the effects of two microfiche-viewing screens on subjects performing a visual task. A high-scintillation viewing screen found to cause visual fatigue, judged by subjects' ratings, produced significantly greater general fatigue than a low-scintillation screen that produced less visual fatigue. The author reported that performance on the visual task was unaffected by the levels of scintillation or fatigue.

Storm and Gray (39) used subjectively reported measures of fatigue, sleep, and biochemical analyses of urine specimens to evaluate the 24-hour work schedule adopted by the U.S. Air Force for Minuteman missile operations in 1977. The authors noted that scores from the Subjective Fatigue Checkcard could be used to discriminate between normal circadian rhythms on off-duty days and rhythms indicating depressed amplitudes and lower overall scores (greater fatigue) during alerts. Significant shifts, noted in circadian rhythms of urinary metabolites and corticosteroids, supported the findings in the subjective measures.

It is apparent that Pearson's Feeling-Tone Checklist and its generic descendent, the Subjective Fatigue Checkcard, have been useful in measuring

fatigue associated with various tasks, conditions, operations, and environments. Because the Subjective Fatigue Checkcard is made up of only 10 statements, it takes a minimum of time to complete and is thus ideal for field studies. In spite of its brevity it appears to describe an activation continuum similar to Thayer's Activation-Deactivation Adjective Checklist which contains more items, takes longer to complete, and does not provide for intensity of feelings. This similarity adds to the evidence that the checkcard provides an integrated, composite description of psychophysiological states.

The Study

The nature of the questions to be answered in this test of the 48-hour work schedule and the nature of the test environment itself suggested the use of subjective measures to assess the effects of the new schedule on crew members. The length of the study, inability to commit experimenters to the test for the entire 90-day period, large number of subjects, and distances of the work centers from the base required that the simplest possible methods of gathering data should be used.

For the purposes of this study, scaled subjective reports of fatigue from the Subjective Fatigue Checkcard were to be used to provide a composite description of the overall physiological and psychological state of respondents at given times. That state was assumed to be the result of experiences since the previous response. This measure was expected to integrate, in a single subjective report, feelings of boredom, physical or mental exhaustion, tension or discomfort, morale, and general health and well being.

To assist the 321st Strategic Missile Wing in assessing the feasibility of implementing 48-hour alerts for missile crews, systematic subjective measures were proposed to provide daily information relevant to crew members' levels of fatigue, quantity and quality of their sleep during alerts and at home, and additional information about their actual activities while on alert. Specific comparisons of reports of fatigue before alerts and immediately after alerts were planned between crew members working the 48-hour work schedule and those working the standard 24-hour schedule. Comparisons of the subjective feelings of fatigue immediately following return to base would be accomplished for both schedules as an indication of safety in driving. Finally, changes in quantity of sleep and levels of fatigue from their pre-alert values would be compared for the two work schedules following one night of recovery at home. These measures after recovery would provide an indication of the psychophysiological costs associated with the work schedules in a fashion similar to the methods suggested by Cameron (7) and Hartman (20).

The primary concern of this test of the 48-hour work schedule was to assure that the stress and strain experienced by crew members working 48-hour alerts were no more difficult or taxing than those experienced by crew members working 24-hour alerts, the accepted standard work schedule for missile operations.

Results of pilot studies of 48-hour work schedules and related operational tests suggested that null hypotheses predicting no differences in all proposed comparisons of the work schedules would not be rejected.

METHOD

Subjects

The subjects were 160 male officers of the nearly 200 Minuteman missile crew members assigned to the 321st Strategic Missile Wing. The activities which they performed during this test were their normal duties.

The two squadrons were chosen for the test because the overall distances of their assigned control centers from the main base were comparable. The control centers assigned to the Wing's other squadron were much closer to the base overall, so it was not included in the test.

Crew members were assigned to their particular squadrons at the time they were initially assigned to the Wing. For purposes of this test, their assignments to squadrons, and therefore to test or control groups, were assumed to be random, because the initial assignments of personnel to the squadrons were not related to factors being evaluated in this test.

Test Group--Sixty-one crew members, who worked 48-hour alerts at the 446th Strategic Missile Squadron's five control centers during the test period, made up the test group.

The test group had a mean experience level of 20 months of missile crew duty and a mean age of 27 years; 63% were married.

Control Group--Ninety-nine officers, who worked 24-hour alerts at the 448th Strategic Missile Squadron's five control centers during the test, made up the control group. The control group had a mean experience level of 20 months and a mean age of 26 years; 55% were married.

Subjects participated at either of two levels of involvement.

1. The crew members normally assigned to either the test or control squadron completed response materials daily throughout the 3-month test. Procedures and requirements described will always apply to these subjects.

2. Crew members who were normally assigned to the 447th Strategic Missile Squadron, the Wing's third squadron, or to the training or evaluation divisions, but who worked some alerts at either the test or control squadrons' control centers during the test period, completed response materials only for those alerts.

Response Materials

Subjective Fatigue Checkcard--This standard form (SAM Form 136) of the USAF School of Aerospace Medicine is a scaled, 10-item questionnaire that uses a 3-choice, Likert response scale for recording a subject's feelings of fatigue. The checkcard yields a score ranging from 0 to 20, with lower scores indicating greater fatigue (see Appendix A, Figure A-1).

Sleep Survey--This standard form (SAM Form 154) of the USAF School of Aerospace Medicine provides a log for recording the previous 24 hours' sleep, and asks questions about ease of going to sleep, quality of sleep, and need for more sleep (see Appendix A, Figure A-2).

Alert Activities Questionnaire--This questionnaire was designed specifically for this study and was used to obtain information about activities during the alert. Specific areas of interest were disruptions of planned sleep schedules and a three-category rating of the intensity of workload for each 24-hour interval (see Appendix A, Figure A-3).

Other Test Procedures

The crew members designated for this study participated in several activities associated with the test that were not part of the criterion measures but may have affected responses to the subjectively reported measures in some way.

Profile of Mood States (POMS)--The POMS is an inventory of 65 items, each rated on a 5-point scale, a product of Educational and Industrial Testing Service, San Diego, California. This inventory identifies six mood factors, and provides tentative norms for scoring and interpreting responses from psychiatric outpatients or normal subjects.

Urine Specimens--Urine specimens were obtained from all crews working alerts during 3 specific periods of the study (days 1-10, 36-47, and 80-90). Each crew member provided a specimen before and after a specified alert tour during each urine collection phase. The maximum number of specimens any crew member was required to provide was 6, i.e., a specimen before and after each of the 3 alert tours.

Procedures

Distribution of Materials--Response materials were provided to the crew members in sealable envelopes with printed instructions to assure that they were aware of times and procedures for completing the forms. They were briefed on the nature of the study and their related responsibilities when they first received materials and on a recurring basis at reissue times. Individual packages were given to them at predeparture briefings before each alert to assure they had materials with them. Additional materials were packaged for off-duty reporting, and extra copies of all forms were available at the squadrons and launch control centers.

Reporting Procedures--The four following schedules for reporting subjective measures were continued throughout the test:

1. Crew members in the test and control groups completed a subjective fatigue checkcard at 1200 and 1800 every day of the 90-day test whether they were on alert, performing other duties, or off duty. While on alert, crew members who were awake at 0400 completed an additional checkcard at that time. Upon returning to the base after each alert the crew members completed a subjective fatigue checkcard regardless of the time.

2. Each day at 1200 the crew members recorded the duration and quality of sleep acquired in the previous 24 hours by appropriately marking a sleep survey.

3. At 1200 following each 24-hour interval of alert duty, the crew members completed an alert activities questionnaire; i.e., members of the control group (24-hour schedule) completed one of these forms for an alert whereas members of the 48-hour test group completed two during an alert.

4. The POMS inventory forms were packaged with the other reporting materials. The crew members completed one at about 1200 every Wednesday during the test regardless of their duty status that day.

Collection of Materials--Sealed envelopes were returned to squadron representatives, debriefing officers, or the Wing project officer. The project officer shipped the sealed envelopes unopened to the USAF School of Aerospace Medicine. This procedure assured the confidentiality of responses. During periods of the study in which urine specimens were collected, crew members collected their own specimens during crew changeover at the launch control centers at approximately 1200. The specimens were collected in small plastic bottles containing dilute hydrochloric acid as a preservative and were brought to the base in an insulated container by the returning crews.

Data Reduction and Analysis

When response materials were received in sealed envelopes at the USAF School of Aerospace Medicine, they were scored by hand and the data were coded for computer storage with identifying information. After all materials were received, scored, and stored following the end of the test, the entire work schedule for the 90 days was re-created from daily alert orders and coded for computer storage to match it with crew members' subjective data. In this way the date and location of all alerts and all responses reported during alerts were available for analysis. A subset of the data was selected from those associated with alerts. This subset represents the data of greatest interest for the initial analysis of this field test.

Variables reported at 1200 the first day, the return day, and 1 day following the end of either 24- or 48-hour alerts were identified as "pre-alert," "end-alert," and "next-day" values respectively. The subjective fatigue response reported on base following return from alert was identified as the "post-return" value. All these values for subjective fatigue, sleep, and alert activities were identified with specific alerts for each crew member.

A subset of these data from alerts was taken by aligning every other alert of the crew members in the 24-hour group with every alert of the crew members in the 48-hour group to a maximum of 10 for any crew member. These were numbered consecutively from the beginning of the test. These "alert numbers" provide information about "alert history," i.e., the number of alerts the crew member has worked since the beginning of the study. To obtain the actual number of alerts worked by members of the control group, it is necessary to multiply these alert numbers by 2.

For each crew member the pre-alert value for fatigue was subtracted from the end-alert and next-day values to produce a "difference" or "change" score. The pre-alert quantity of sleep was subtracted from the next-day quantity. These difference values reduced the variability between subjects, because they reflect only the change within each subject's reports at 1200 before, after, and following recovery from a given alert. The post-return score for fatigue was an event-related rather than time-related value because it was reported when the crew member returned to the base. The post-return values were treated as absolute scores in the analyses, because there appeared to be no event to compare them to for producing change scores or difference values. The difference variables used in the analyses were renamed "alert" and "recovery" for convenience in describing them and the post-return score was simply called "return."

Using difference scores for variables further reduced the available data set, because only alerts for which a subject provided both pre- and end-alert or pre-alert and next-day reports were included in the analyses of variance.

All statistical analyses were performed on the San Antonio Data Service Center computer using programs from the "Statistical Analysis System" (3).

Four independent analyses of variance were performed on each of the dependent variables (alert, return, and recovery fatigue; and recovery sleep). The first three of these included tests for the effects of uncontrolled variables about which information was available. These included: (a) alert activities, (b) alert history, and (c) control centers. The fourth set of analyses were to test the overall effect of schedule, but it was necessary to obtain the results of the preceding analyses prior to performing the fourth set. In this way the proper design for including factors would be known.

The error term for computing F ratios in these analyses was based on the variance between repeated measures, a within-subject variance. For this reason between-group measures were best tested in the final model in which a between-subjects error term was estimated by a linear combination of the mean squares of subjects and repetitions within subjects. A less sensitive test of the effect of the work schedules on the groups was possible in the nested analyses for the main effect of control centers.

Analyses of variance were used to determine if there were interactions between the dependent variables and the alert history of a crew member; i.e., "were there carryover effects that increased or decreased as the number of alerts worked by a crew member increased?" This question was tested along with the tests for interactions between the schedule and subjectively reported levels of alert activities which were uncontrolled variables in the study. The effect of alert history was also tested in a set of analyses that excluded the alert activities variables. It was necessary to consider the results of both sets of analyses together to assure that including alert activities did not mask real effects. Tables 1 and 2 illustrate the models used for each of these sets of analyses.

TABLE 1. MODEL OF ANALYSES OF VARIANCE FOR INTERACTION BETWEEN WORK SCHEDULE AND ALERT HISTORY AND ALERT ACTIVITIES (SLEEP DISRUPTION AND WORKLOAD)

Source of variation	df
Between crew members (Ss)	
Work schedule (G)	1
Crew members (Ss)	-[a]
Within crew members (Ss)	
Control centers (CC)	8
Alert history (T)	9
G x T	9
Sleep disruption (SL)	1
G x SL	1
Workload (WL)	2
G x WL	2
Error	-[a]

[a]Degrees of freedom differed in each analysis because there were incomplete data.

TABLE 2. MODEL OF ANALYSES OF VARIANCE FOR INTERACTION BETWEEN WORK SCHEDULE AND ALERT HISTORY

Source of variation	df
Between crew members (Ss)	
Work schedule (G)	1
Crew members (Ss)	-[a]
Within crew members (Ss)	
Control centers (CC)	8
Alert history (T)	9
G x T	9
Error	-[a]

[a]Degrees of freedom differed in each analysis because there were incomplete data.

To provide tests of the effect of control centers, nested analyses using a further reduced subset of the alert data were performed. In these analyses only dependent variables reported by crew members at one control center, where they worked most often, were included in the data. Thus crew members were nested within control centers and a linear combination of the mean squares between crew members and repetitions within crew members provided the proper error term for testing the effect of control centers. These nested analyses also provided the most sensitive comparison of the between-subject-within-group and the repetitions-within-subjects variabilities. Further, the results of these nested analyses of the effect of control centers could determine which variables should be included in the analyses for effect of the work schedule. Table 3 summarizes the model used for the nested analyses for control centers.

TABLE 3. MODEL OF ANALYSES OF VARIANCE FOR THE MAIN EFFECT OF CONTROL CENTERS

Source of variation	df
Between work schedule (G)	
Work schedule (G)	1
Error	_[a]
Within work schedule (G)	
Control centers (CC)	8
Error	_[a]
Crew members (Ss)	_[b]
Error	_[b]

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Degrees of freedom differed in each analysis because there were incomplete data.

The design of the model described in Table 4 was based on the previously described analyses. In this model an estimated mean square is used to calculate the proper F ratios to test for the main effect of work schedule. However, the elimination of the previously included independent variables makes these analyses the most sensitive tests of the effect of the work schedule that are possible from these data. The subset of data used in the two analyses for interactions was used for these analyses, but in these data control centers, alert history, and alert activities were ignored to provide the best overall test.

TABLE 4. MODEL OF ANALYSES OF VARIANCE FOR THE MAIN EFFECT OF WORK SCHEDULE

Source of variation	df
Between work schedule (G)	
Work schedule (G)	1
Error	_[a]
Within work schedule (G)	
Crew members (Ss)	_[b]
Error	_[b]

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Degrees of freedom differed in each analysis because there were incomplete data.

RESULTS

The analyses of variance indicated that there were no significant interactions between alert activities (disrupted sleep and different levels of workload) and the two work schedule conditions for the dependent variables of subjective fatigue (alert, return, and recovery reports) or quantity of sleep during recovery. Also, no significant interactions between the schedule worked and the alert history were found for the dependent variables of fatigue and sleep.

The nested analysis showed that the effect and control centers within work schedules approached a significant level for the dependent variable of subjective fatigue reported after return from alert, $F(8,82) = 1.97$, $p = .061$. In nested analyses of the other dependent variables the effects of control centers did not approach significant levels. The previously described insensitive tests of the effects of work schedule in these nested analyses indicated no significant differences in the levels of subjective fatigue reported. However, in the analysis of the effects of work schedule on recovery sleep, a significant difference was found between the test and control group, $F(1,7) = 14.25$, $p < .01$. Tests of the variance of subjects within groups were significant at probabilities of less than .001 for all dependent variables in these nested analyses.

A final test of analyses was performed in which all the variances were pooled by summation of the previously analyzed independent variables. In the analysis of variance of the effect of work schedule on the dependent variable

of recovery sleep, a significant difference between the two groups was found, $F(1,72) = 10.19$, $p < .01$. In the remainder of this set of analyses of the dependent variables of subjective fatigue no significant differences were found. Tables 5 - 20 provide summaries of all the analyses of variance that were performed.

TABLE 5. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY AND ALERT ACTIVITIES (SLEEP DISRUPTION AND WORKLOAD): ALERT FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	108		
Work schedule (G)	1	-[a]	
Crew members (Ss)	107	-[a]	
Within crew members (Ss)	385		
Control centers (CC)	8	8.23	.58
Alert history (T)	9	6.43	.34
G x T	9	12.73	.89
Sleep disruption (SL)	1	0.00	.00
G x SL	1	8.40	.59
Workload (WL)	2	11.76	.82
G x WL	2	7.97	.56
Error	353	14.27	
Total	493		

[a]Computer program did not provide correct mean squares because of the method of analyses of these data.

TABLE 6. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY AND ALERT ACTIVITIES (SLEEP DISRUPTION AND WORKLOAD): RETURN FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	111		
Work schedule (G)	1	-[a]	
Crew members (Ss)	110	-[a]	
Within crew members (Ss)	327		
Control centers (CC)	8	5.27	.73
Alert history (T)	9	13.62	1.88
G x T	9	9.41	1.30
Sleep disruption (SL)	1	23.84	3.29
G x SL	1	5.33	.73
Workload (WL)	2	25.01	3.45[b]
G x WL	2	7.08	.97
Error	295	7.26	
Total	438		

[a]Computer program did not provide correct mean squares because of the method of analysis of these data.

[b]$p < .05$

TABLE 7. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY AND ALERT ACTIVITIES (SLEEP DISRUPTION AND WORKLOAD): RECOVERY FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	82		
Work schedule (G)	1	-[a]	
Crew members (Ss)	81	-[a]	
Within crew members (Ss)	345		
Control centers (CC)	8	9.17	.62
Alert history (T)	9	11.84	.79
G x T	9	6.00	.40
Sleep disruption (SL)	1	2.64	.18
G x SL	1	3.62	.24
Workload (WL)	2	23.67	1.59
G x WL	2	6.30	.42
Error	313	14.90	
Total	427		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

TABLE 8. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY AND ALERT ACTIVITIES (SLEEP DISRUPTION AND WORKLOAD): RECOVERY SLEEP

Source of variation	df	MS	F
Between crew members (Ss)	85		
Work schedule (G)	1	-[a]	
Crew members (Ss)	84	-[a]	
Within crew members (Ss)	371		
Control centers (CC)	8	5.48	1.53
Alert history (T)	9	5.73	1.59
G x T	9	3.28	.91
Sleep disruption (SL)	1	2.23	.62
G x SL	1	.11	.03
Workload (WL)	2	2.83	.53
G x WL	2	2.60	.72
Error	339	14.90	
Total	456		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

TABLE 9. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY: ALERT FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	113		
Work schedule (G)	1	-[a]	
Crew members (Ss)	112	-[a]	
Within crew members (Ss)	434		
Control centers (CC)	8	7.47	.53
Alert history (T)	9	5.79	.41
G x T	9	10.80	.77
Error	408	14.08	
Total	547		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

TABLE 10. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTION BETWEEN WORK SCHEDULE AND ALERT HISTORY: RETURN FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	113		
Work schedule (G)	1	-[a]	
Crew members (Ss)	112	-[a]	
Within crew members (Ss)	352		
Control centers (CC)	8	7.65	1.0
Alert history (T)	9	18.67	2.43[b]
G x T	9	6.32	.82
Error	326	7.68	
Total	465		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

[b]$p < .05$

TABLE 11. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTION BETWEEN WORK SCHEDULE AND ALERT HISTORY: RECOVERY FATIGUE

Source of variation	df	MS	F
Between crew members (Ss)	85		
Work schedule (G)	1	-[a]	
Crew members (Ss)	84	-[a]	
Within crew members (Ss)	385		
Control centers (CC)	8	9.80	.67
Alert history (T)	9	14.12	.97
G x T	9	8.23	.57
Error	359	14.53	
Total	470		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

TABLE 12. SUMMARY OF THE ANALYSIS OF VARIANCE FOR INTERACTIONS BETWEEN WORK SCHEDULE AND ALERT HISTORY: RECOVERY SLEEP

Source of variation	df	MS	F
Between crew members (Ss)	90		
Work schedule (G)	1	-[a]	
Crew members (Ss)	89	-[a]	
Within crew members (Ss)	422		
Control centers (CC)	8	5.49	1.53
Alert history (T)	9	5.17	2.00[b]
G x T	9	5.41	1.51
Error	396	3.58	
Total	512		

[a]The computer program did not provide correct mean squares because of the method of analysis of these data.

[b]$p < .05$

TABLE 13. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF CONTROL CENTERS: ALERT FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	1.99	.09
Error	6.8[a]	23.36[b]	
Within work schedule (G)	507		
Control centers (CC)	8	24.77	.44
Error	79.9[a]	56.45[b]	
Crew members (Ss)	97	43.75	3.24[c]
Error	402	13.51	
Total	508		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 14. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF CONTROL CENTERS: RETURN FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	4.71	.06
Error	7.5[a]	84.10[b]	
Within work schedule (G)	418		
Control centers (CC)	8	81.22	1.97
Error	82.4[a]	41.29[b]	
Crew members (Ss)	102	28.51	3.80[c]
Error	308	7.51	
Total	419		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 15. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF CONTROL CENTERS: RECOVERY FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	1.24	.16
Error	5.2[a]	7.70[b]	
Within work schedule (G)	460		
Control centers (CC)	8	9.03	.28
Error	54.8[a]	32.23[b]	
Crew members (Ss)	73	26.76	2.12[c]
Error	379	12.65	
Total	461		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 16. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF CONTROL CENTERS: RECOVERY SLEEP

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	75.69	14.25[c]
Error	6.5[a]	5.31[b]	
Within work schedule (G)	507		
Control centers (CC)	8	5.51	.56
Error	60.1[a]	9.83[b]	
Crew members (Ss)	78	8.09	2.29[d]
Error	421	3.54	
Total	508		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c] $p < .05$

[d] $p < .001$

TABLE 17. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF WORK SCHEDULE: ALERT FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	5.96	.13
Error	92.5[a]	46.63[b]	
Within work schedule	546		
Crew members (Ss)	112	38.36	2.79[c]
Error	434	13.75	
Total	547		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 18. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF WORK SCHEDULE: RETURN FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	38.64	.84
Error	97.9[a]	45.97[b]	
Within work schedule (G)	474		
Crew members (Ss)	112	35.10	4.43[c]
Error	352	7.92	
Total	475		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 19. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF WORK SCHEDULE: RECOVERY FATIGUE

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	1.37	.05
Error	63.1[a]	26.53[b]	
Within work schedule (G)	469		
Crew members (Ss)	84	23.87	1.66[c]
Error	385	14.37	
Total	470		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .001$

TABLE 20. SUMMARY OF THE ANALYSIS OF VARIANCE FOR THE MAIN EFFECT OF WORK SCHEDULE: RECOVERY SLEEP

Source of variation	df	MS	F
Between work schedule (G)	1		
Work schedule (G)	1	86.68	10.19[c]
Error	72.0[a]	8.51[b]	
Within work schedule (G)	511		
Crew members (Ss)	89	7.54	2.03[d]
Error	422	3.71	
Total	512		

[a]Degrees of freedom were approximated from the mean squares used in linear combination to estimate the error term.

[b]Estimated by linear combination of mean squares to obtain correct expected value.

[c]$p < .01$

[d]$p < .001$

DISCUSSION

A systematic difference in the quantity of sleep acquired, on the first night at home after working alerts, was the single factor that distinguished crew members on the 48-hour work schedule from those on a standard 24-hour schedule. This finding was significant in an insensitive analysis with few degrees of freedom and in a considerably more sensitive analysis with a large number of degrees of freedom. The fact that it was found in both analyses indicates that it probably was not a spurious result.

Hartman (20) demonstrated that increases in quantity of sleep of aircrews following the end of strenuous missions were an indication of the effort required to fly the missions. Cameron (7) had recommended measuring the time required to recover to a normal state, determined by undefined physiological criteria, would be the best way to compare how much effort was required by conditions expected to cause fatigue. In both approaches the duration of the events or conditions being compared is not important, but the times to recover are. If recovery times differ, it is likely the effort required by the two conditions differs also.

In this study subjective reports of fatigue were collected as a composite description of crew members' overall psychophysiological state. They attempted to measure the "phenomenological awareness of bodily functioning" described by Thayer (44). It is important to note that these reports did not differ significantly between the two groups at any of the times measured, particularly the report following recovery sleep.

This fact must be considered in judging the feasibility of the schedule. Crew members who worked 48 hours appeared to require more sleep after their alerts than those completing 24-hour alerts, but an average of slightly more than an extra half hour of sleep was sufficient for them to match the subjective feelings of members of the control group who worked the standard alert. The extra 30 minutes of sleep appears to be a small price to pay in order to obtain the economic and other gains offered by the 48-hour work schedule.

Is this additional sleep by the members of the test group really "required"? There are other reasons why the crew members of the 48-hour schedule may have slept longer. The regulations governing missile crews' rest and recuperation following alerts (45) require that for each hour in the underground control center the crew member is "guaranteed" one-half hour free of duty. This time begins when he reports back to base after the alert. Under normal conditions crew members do not have to work alerts the day following their return day, but they are scheduled for other duties. The 48-hour schedule guaranteed that crew members returning from a 48-hour alert would not be scheduled for any duties until the afternoon of the day after returning to base. In most cases they were completely free of duty on that day also. Therefore the crew members in the test group were officially provided an opportunity to obtain a greater amount of sleep than the crew members who worked standard alerts who may have had to arise to perform official duties on the day after their return to base.

The small difference in sleep between the groups during recovery was not associated with any measurable difference in feelings of fatigue at the report following that sleep. Because the increased sleep may have resulted from greater opportunity for the test group to sleep, this finding is considered insufficient for declaring the 48-hour work schedule infeasible.

The measures in this study relied almost entirely on subjective reports by the crew members. It is important, therefore, to consider how well these data reflect what actually occurred during the study and what threats to the validity of the results or to generalizations from the results may exist.

A model crew member in either the test or control group, who remained in the test through the entire 90 days and worked the maximum number of alerts, would have completed over 350 subjective reports. In fact no crew member completed all reporting requirements, and many completed a very low ratio of the required reports. Because the reasons for subjects' not completing specific measures may have been related to the variables of interest in the test, low ratios of completion are a threat to the validity of the results.

Initial briefings encouraged full participation throughout the study by all participants. When it became evident that some crew members were not completing all reports, greater emphasis on the importance of complete data

was stressed in briefings. It would have been possible to use punitive or at least threatening procedures against individual crew members who cooperated very little, but such methods were not used because employing them could have caused deleterious effects on the "quality" of other subjects' reports. Instead the briefings were continued throughout the test to encourage and motivate the crew members to complete the required forms.

An initial scan of the data showed that crew members completed forms immediately before and during alerts much more often than they did when off duty. Because the subset of the data used for the analyses of variance was drawn primarily from alerts (next-day responses were usually reported when off duty), it represented the most complete data and appeared to be the most valid for analysis.

Still, within these data from alerts it was possible that failures to complete reports were the result of fatigue. In that case these missing reports, by their absence, could bias the results in favor of finding no effect of fatigue. Because the primary interest of the study was to determine if the test group was more fatigued than the control group, rather than to determine the "true" fatigue levels experienced by the crew members, it was possible to estimate the effect or direction of such bias on the data by comparing the ratios of completed forms between the groups. (See Appendix B for the results of these comparisons.) The test group completed significantly more of the required reports than the control group did. With this factor taken into account, if fatigue were in fact the cause of the failure to respond, the test group was less likely than the control group to have had results biased because of missing data. Two possible reasons why no significant differences between the groups were found in reports of fatigue are: (a) either fatigue was not the cause of the failure to report, or (b) the control group was in fact more fatigued than the test group, but did not complete the forms that would have shown this result. Neither case supports an argument against the feasibility of implementing the 48-hour work schedule.

The measures in this study were recorded daily throughout a 90-day period. It was hoped that all possible situations that would provide a thorough test of the feasibility of the 48-hour schedule would occur randomly throughout this period. Previous study of missile operations indicated, however, that systematic effects might be present within one group that might bias the results in a particular direction.

More important, if the crew members working the 48-hour schedule were differentially affected by systematic effects such as workload, it may have been infeasible to implement the schedule. The investigation of systematic effects is discussed here.

The 2-member crews worked most often at 1 control center of the 5 assigned to their squadron. Although the control centers were assumed to be comparable between the 2 squadrons, differences among them could have affected the dependent variables measured. Two differences that were not subject to change during the study were the distances from the main base and the on-site equipment configurations. The round-trip mileage affected travel time and the overall length of an alert tour. The control centers that were configured as squadron command posts had additional communications equipment and more

responsibilities for command and control than other control centers. The squadron command post configuration would be expected to affect workload for the crew members. Table 21 indicates the control center designations, round-trip distances, and configurations.

TABLE 21. CONTROL CENTER DIFFERENCES: DISTANCES AND EQUIPMENT

Designation	Mileage[a]	Configuration[b]
	Test squadron	
A	237	SCP
B	212	
C	140	SCP
D	168	
E	200	
	Control squadron	
K	80	SCP
L	146	
M	149	
N	181	SCP
O	142	

[a]Round-trip distance.

[b]SCP = Squadron Command Post.

Alert tours at the 10 control centers could have differed from each other in several ways in addition to the distances from the main base and in equipment configurations. Four examples are the following:

1. Maintenance activities at the launch control center or any of the launch facilities controlled by it require the crew to participate by controlling access, briefing procedures for maintenance, and monitoring some activities visually or by audio communications.

2. Communications from outside sources, especially during local or worldwide exercises, require the crews to take some action. Active communications add to the ambient noise levels as well as the workload.

3. Security alarms triggered at any of the missile launch facilities require the crews to coordinate communications between security teams and the main base until the situation is resolved.

4. Malfunctions of a control center's equipment could cause special problems such as increased noise level, vibration, or extremes in environmental temperature. The crew could have to use extraordinary procedures to compensate for the malfunctioning equipment.

These are a few of the events that could make one alert different from another. Such problems could occur randomly, affecting 1 or 2 control centers and their crews for several hours by increasing their workload and/or disrupting their sleep during their alert tour. But there could be consistent problems at 1 or several control centers for several days or even weeks that could affect crews who worked there every time they were on alert.

Because each control center and each alert that is worked there could differ from other control centers and other alerts in a nonindependent manner, tests for the effects of control centers and activities while on alert were added to the analyses of variance.

Because no interactions of the work schedule and the reports of sleep disruption or levels of workload were found to be significant for the subjective reports of fatigue or the quantities of sleep during recovery, it appears that no systematic effects of this nature were operating during the study. However, these measures of sleep disruption and workload were reported subjectively, just as the fatigue and sleep variables were. It was very likely that these values could be highly correlated with the dependent variables, because of this similarity in measurement. Different methods of measuring these data may have provided different results.

Because no effects of control centers on these dependent variables were found, it appears that control centers were comparable within the groups.

In summary the variables measured in this study did not indicate that crew members working either a 48-hour schedule or the standard work schedule were differentially affected by activities that occurred during alerts or by the characteristics or locations of the specific control centers where they worked.

Another systematic effect that could have influenced the results was the possibility of differential changes over the course of the study. Some changes in subjective feelings might be expected in both groups. Winter weather, which became progressively worse throughout the study, could be expected to be an influence. But relevant to the schedules, little change was expected in the crew members working 24-hour alerts, because they were doing essentially the same work they had done during the previous 90 days. In contrast the members of the test group were starting new procedures, and there was some probability that the effects of these new procedures might differ

throughout the measurement period. The initial effect could be an increase in reported levels of fatigue as crew members learned how to manage their time and energy and adjusted to the new work schedule. Later, after this adjustment, their feelings of fatigue might improve. Another possibility was that the initial effects of one or two of the 48-hour alerts would be benign, but some cumulative effect that would only become evident over many alerts would later increase subjective feelings of fatigue or quantity of sleep required during recovery.

The analyses of the effects of alert history were included to deal with the possibility of such cumulative effects. The number associated with each alert provided an indirect indication of the time elapsed since the beginning of the study and the exact history of the alerts a crew member had worked.

The failure to find a significant interaction of the work schedule and the number of alerts worked indicates that there were no differential changes between the test and control groups throughout the course of the study. This finding allowed the pooling of data from all alerts in order to perform the analyses of the main effect of the schedules.

The threats to validity just described are those discussed by several authors as relevant to internal validity (8, 11, 27). These threats to validity may directly affect the results of the study and can often be dealt with in the analyses.

Additional circumstances in the study influenced the external validity, i.e., the ability to generalize the results to other situations. This type of validity is usually dependent on how subjects are selected or sampled in the research.

The crew members in the test and control groups were not assigned randomly. Their initial assignments to the squadron used as test and control groups could be considered random, because such assignments were unrelated to the variables of concern in this study.

However, the selection of test group status for the 446th Strategic Missile Squadron was based on situations that could have influenced the results obtained; it was not a random process. These situations were:

1. The project officer assigned by the 321st Strategic Missile Wing to evaluate the test was a member of the test squadron.

2. The project officer and other members of the test group had participated in the pilot study of the 48-hour work schedule (40).

In addition to the influence that nonrandom selection may have had on the results, the treatment of the test and control groups throughout the study also may have influenced the results obtained.

The treatment of crew members in the test and control groups was not the same. This was particularly the case in activities that the Wing implemented to study the effects of the alert schedule concurrently with the methods used by the USAF School of Aerospace Medicine. In most instances the Wing placed

greater emphasis on the test group's participation than on the control group's. Examples of this emphasis were:

1. Throughout the study crew members in the test group were interviewed by the Wing's project officer, and they completed questionnaires each month regarding their morale, their families' opinions, and other aspects of the test for the Wing's own study of the problem.

2. Throughout the study these crews were required to give a special debriefing after their alerts.

3. All the crews from the test group were evaluated for procedural performance in the missile procedures trainer after returning from one of their alert tours during the 90-day period. These evaluations were observed by the Wing evaluator crews and were similar to an annual proficiency check. Only 10 randomly selected crews from the control group were evaluated by this procedure.

4. The fact that the project officer and the evaluator from the USAF School of Aerospace Medicine worked in offices in the test squadron may have influenced the crew members in that group, because they were more familiar with the evaluators and may have had conversations or interactions with these individuals more often than members of the control group did.

The overall effect of these processes may have caused a positive bias to the subjective reports of the test group's crew members; i.e., their responses would be more favorable to the 48-hour schedule. It is also possible that instead the increased scrutiny was judged negatively by some crew members. No attempt was made to measure directly the effects of these biased treatments of the groups. If these overall effects were sufficient to bias the results in favor of the 48-hour work schedule, it appears that these influences would not make the schedule infeasible within this group. But similar influences and involvement may be necessary to introduce this schedule to crew members who are not familiar with it. The results might have been different if the crew members had perceived that the schedule was being imposed upon them.

To generalize the results of a study such as this to nonstudy situations, i.e., the implementation of the 48-hour work schedule in normal missile operations, it is necessary to consider what influence the procedures in the study may have had on the results. This has been done to some extent in discussing possible differential treatments of the groups. The requirements of completing response materials daily throughout the 90-day period may have affected the crew members in ways that would not be found in normal operations. In addition to the collection of data from which interferences were made in this study, crew members participated in two other types of data collection described under Methods. These were the weekly completion of the POMS inventory and the collection of urine specimens at three phases of the test.

Completing the POMS was very similar to the requirements of the other subjective assessments. It would be expected to influence subjects similarly. Both groups were treated the same, relevant to the POMS.

Hartman (21) has stated that the sampling of biological measures can improve the subjects' interest and belief in a study and can cause them to answer subjective responses more "truthfully." However, an opposite effect may have been the case. Specifically, a U.S. Air Force-wide program of sampling urine to detect drug abuse was being reestablished at all U.S. Air Force bases during the time of this missile test. The negative attitude that has often been expressed toward this drug testing may have carried over to the sampling of urine and other test requirements as well. No attempt was made to measure either of these effects. In either case both groups were treated equally in the sampling of specimens.

A final question asked in a study of this nature, particularly one in which few significant results of the treatments were found, is, "Were the proper variables measured?"

The nearly total use of subjectively reported measures is a likely source of criticism. Critics of the use of subjective reports have focused primarily on the failure of such measures to predict levels of performance in laboratory experiments (4, 37). Some studies of sleep deprivation (28, 29, 32) have shown that sleep-deprived subjects who reported extreme fatigue were still able to maintain adequate performance on psychomotor and cognitive tasks. Other authors have countered these criticisms of subjective measures by contrasting the laboratory environment from which they are drawn with the real-world, work environment. Cameron (7) reported that the motivation to perform well in relatively short-term laboratory settings is likely to cause an experimental subject to overcome feelings of fatigue in order to maintain performance. Such a high level of motivation cannot be assumed to exist in the day-to-day lives of individual workers. Cameron suggested that performance would more closely parallel feelings of fatigue in the working environment. This general argument is supported by an examination of procedures used in sleep deprivation studies (17).

Subjects in these studies received a great deal of social and emotional support from the efforts of experimenters and other subjects to keep them awake even though they experienced feelings of fatigue. The authors explained the maintenance of adequate performance in these studies as a product of social support and motivation. To counter such effects, Gifford and Murawski (17) performed sleep deprivation experiments using isolated individuals and isolated pairs of subjects. In their studies the performance of tasks and the ability to remain awake were both greatly reduced. According to the authors, the circumstances of these experiments more closely simulated real-world sleep deprivation problems than the experiments previously discussed.

The problem of motivation is further illustrated by circumstances within the present study. The missile Wing attempted to measure the effects of working the 48-hour alert schedule by evaluating the missile crews in the missile procedures trainer after they returned from alert. Because these evaluations contained sensitive information, they could not be observed by the USAF School of Aerospace Medicine's investigator. The evaluations were observed only by personnel of the 321st Strategic Missile Wing. The results of the evaluations were:

1. Two of the 10 crews from the control group failed.

2. None of the 25 crews from the test group failed.

Note, these were not individual evaluations; the 2 crew members performed as a team. The evaluators reported that although most crew members passed all evaluations, they appeared more disorganized in their procedures than is normally the case in evaluations. It is important to note that the annual evaluations are never accomplished immediately following alerts.

Did these evaluations answer the question of crew effectiveness? It is certainly true that this information contributes to and supports the other findings, but these evaluations for a single crew member or even a single crew evaluate the effects of only 1 of the 10 or 12 alerts worked during the test. The fact that the only crews who failed were from the group least expected to be fatigued makes interpretation of the results somewhat difficult. Was some subjective factor such as greater motivation in the 48-hour group operating in this situation?

Subjective measures have been found to agree with behavioral and physiological measures in many field studies, and the Subjective Fatigue Checkcard used in this study has been validated as a measure of fatigue resulting from various sources. It appears that sufficient opportunity was provided for disadvantages of the 48-hour schedule to become obvious. The fact that no disadvantages were found makes it reasonable to conclude that the 48-hour work schedule as tested in this study is feasible for Minuteman ICBM operations. As noted earlier, caution should be used in implementing the work schedule to assure circumstances similar to the test environment are employed; i.e., the new work schedule should not be imposed on the crew members. If that is done, results may be different from those found in this evaluation.

ACKNOWLEDGMENTS

Several individuals made significant contributions to this study. Dr. William F. Storm was instrumental in the initial planning and made helpful suggestions throughout the study. Sergeant Robin Chavez and A1C Vabian Paden completed the reduction of the enormous amount of data and assisted in the analysis.

An extremely significant contribution was made by Mr. Richard McNee, Data Sciences Division, whose helpful suggestions, guidance, and thorough expertise in statistical analysis continued from the planning through the final writing of this report.

REFERENCES

1. Air Force Pamphlet 161-18. Flight surgeons guide. Washington, D.C.: Department of the Air Force, Headquarters U.S. Air Force, Dec 1968.

2. Ash, I. Fatigue and its effects upon control. Arch Psychol No. 31 (1914).

3. Barr, A. J., and J. H. Goodnight. A users guide to the statistical analysis system. Raleigh, N.C.: North Carolina State University, Aug 1972.

4. Bartlett, F. Psychological criteria of fatigue. In W. F. Floyd and A. T. Welford (eds.). Symposium on fatigue. London: H. K. Lewis & Co., 1953.

5. Bartley, S. H., and E. Chute. Fatigue and impairment in man. New York: McGraw-Hill, 1947.

6. Browne, R. C. Fatigue, fact or fiction. In W. F. Floyd and A. T. Welford (eds.). Symposium on fatigue. London: H. K. Lewis & Co., 1953.

7. Cameron, C. A theory of fatigue. Ergonomics 16:633-648 (1973).

8. Campbell, D. T., and J. C. Stanley. Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Company, 1966.

9. Cantrell, G. K., R. W. Trimble, and B. O. Hartman. Long-term aircrew effectiveness (A literature survey). SAM-TR-71-4, 1971. (NTIS No. AD-722 417)

10. Colquhoun, W. P. (ed). Biological rhythms and human performance. New York: Academic Press, 1971.

11. Cook, T. D., and D. T. Campbell. Quasi-experimentation: Design & analysis issues for field settings. Chicago: Rand McNally College Publishing Company, 1979.

12. Crawford, A. Fatigue and driving. Ergonomics 4:143-154 (1961).

13. Cushman, W. H. Visual search using microfiche: Effects of microfiche reader scintillation on visual fatigue, general fatigue, and performance. In E. J. Baise and J. M. Miller (eds.). Proceedings of the Human Factors Society, 22nd Annual Meeting, Detroit, Mich., 1978.

14. Duffy, E. Activation and behavior. New York: John Wiley and Sons, 1962.

15. Dukes-Dobos, F. N. Fatigue from the point of urinary metabolites. In K. Hashimoto, K. Kogi, and E. Grandjean (eds.). Methodology in human fatigue assessment. London: Taylor & Francis Ltd., 1971.

16. Eason, R. C., and L. M. Dudley. Physiological and behavioral indicants of activation. Psychophysiology 7:223-232 (1971).

17. Gifford, S., and H. L. Murawski. Minimal sleep-deprivation alone and in small groups: Effects on ego-functioning and 24-hour body temperature and adrenocortical patterns. Symposium on Medical Aspects of Stress in the Military Climate. Walter Reed Army Institute of Research, 1964.

18. Grandjean, E. Fatigue: Its physiological and psychological significance. Ergonomics 11:427-436 (1968).

19. Harris, D. A., et al. Oral temperature in relation to inflight work/rest schedules. Aerosp Med 41:723-727 (1970).

20. Hartman, B. O. Field study of transport aircrew workload and rest. Aerosp Med 42:817-821 (1971).

21. Hartman, B. O. USAF School of Aerospace Medicine, Brooks AFB, Tex. Personal communication, Sep 1976.

22. Hartman, B. O. Management of irregular rest and activity. Sleep, wakefulness, and circadian rhythm (AGARD Lecture Series, No. 105). Neuilly Sur Seine, France: Advisory Group for Aerospace Research & Development, North Atlantic Treaty Organization, Sep 1979. (NTIS No. AD-A075240)

23. Hartman, B. O., et al. Operational aspects of variations in alertness (AGARDograph No. 189). Neuilly Sur Seine, France: Advisory Group for Aerospace Research & Development, North Atlantic Treaty Organization, Aug 1974. (NTIS No. AD-786 123/OGI)

24. Hartmann, E. The functions of sleep. New Haven: Yale University Press, 1973.

25. Holden, L. Request for consultant assistance. Formal letter. Grand Forks AFB, N.D.: Headquarters, 321st Strategic Missile Wing (SAC), Aug 14, 1978.

26. Innes, L. G. A subjective assessment of fatigue in transport aircrew. Rest and Activity Cycles for the Maintenance of Personnel Concerned with Military Flight Operations (AGARD Conference Proceedings No. 74). Neuilly Sur Seine, France: Advisory Group for Aerospace Research & Development, North Atlantic Treaty Organization, Nov 1970. (NTIS No. AD-717 265).

27. Isaac, S., and W. B. Michael. Handbook in research and evaluation. San Diego: Educational and Industrial Testing Service, 1971.

28. Johnson, L. C. Sleep disturbances and performance. Sleep, wakefulness, and circadian rhythm (AGARD Lecture Series No. 105). Neuilly Sur Seine, France: Advisory Group for Aerospace Research & Development, North Atlantic Treaty Organization, Sep 1979. (NTIS No. AD-A075240)

29. Lubin, A. Performance under sleep loss and fatigue. In S. S. Kety, E. V. Evarts, and H. L. Williams (eds.). Sleep and altered states of consciousness. Baltimore: Williams & Wilkins, 1967.

30. McFarland, R. A. Fatigue in industry: Understanding fatigue in modern life. Ergonomics 14:1-10 (1971).

31. Mohler, S. R. Physiological index as an aid in developing airline pilot scheduling patterns. Aviat Space Environ Med 47:238-247 (1976).

32. Murray, E. J., H. L. Williams, and A. Lubin. Body temperature and psychological ratings during sleep deprivation. J Exp Psychol 56:271-273 (1958).

33. Muscio, B. Is a fatigue test possible? Br J Psychol 12:31-46 (1921).

34. Muscio, B. Feeling-tone in industry. Br J Psychol 12:150-162 (1921).

35. Pearson, R. G. Scale analysis of a fatigue checklist. J Appl Psychol 41:186-191 (1957).

36. Pearson, R. G., and G. E. Byars. The development and validation of a checklist for measuring subjective fatigue. SAM-TR-56-115. Randolph AFB, Tex: USAF School of Aviation Medicine, Apr 1956.

37. Poffenberger, A. T. The effects of continuous mental work upon output and feelings. J Appl Psychol 12:459-467 (1928).

38. Shotts, M. 48-hour Minuteman alert test. Formal letter. March AFB, Calif.: Headquarters, 15th Air Force (SAC), Jul 31, 1978.

39. Storm, W. F., and S. F. Gray. Minuteman missile crew fatigue and 24-hour alerts. SAM-TR-78-19, May 1978. (NTIS No. AD-A056561)

40. 321st Strategic Missile Wing. 48-hour alert test program final report. Grand Forks AFB, N.D., Jun 1978.

41. 321st Strategic Missile Wing. Extended 48-hour Minuteman alert test: Operations Plan 411-78. Grand Forks AFB, N.D.: Headquarters, 321st Strategic Missile Wing, Aug 1978.

42. 390th Strategic Missile Wing. Titan II 48-hour alert test program study. Davis-Monthan AFB, Ariz., Apr 1977.

43. Sullivan, E. T. Mood in relation to performance. Arch Psychol No. 53 (1922).

44. Thayer, R. Measurement of activation through self report. Psychol Rep 20:663-687 (1967).

45. USAF Strategic Air Command Regulation 55-43, SAC Alert Procedures. Washington, D.C.: Department of the Air Force, Headquarters, U.S. Air Force, Oct 1979.

46. Webb, W. B., and H. W. Agnew. Effects of performance of high and low energy-expenditure during sleep deprivation. Percep Mot Skills 37:511-514 (1973).

47. Welford, A. T. Stress and performance. Ergonomics 16:567-580 (1973).

48. Weybrew, B. B. Patterns of psychophysiological response to military stress. In M. H. Appley and R. Trumbull (eds.). Psychological stress: Issues in research. New York: Appleton-Century-Crofts, 1967.

49. Wolf, G. Construct validation of measures of three kinds of experimental fatigue. Percept Mot Skills 24:1067-1076 (1967).

APPENDIX A

INSTRUMENTS FOR SUBJECTIVE REPORTS

NAME AND GRADE			TIME/DATE
INSTRUCTIONS. Make one and only one (√) for each of the ten items. Think carefully about how you feel RIGHT NOW.			
STATEMENT	BETTER THAN	SAME AS	WORSE THAN
1. VERY LIVELY			
2. EXTREMELY TIRED			
3. QUITE FRESH			
4. SLIGHTLY POOPED			
5. EXTREMELY PEPPY			
6. SOMEWHAT FRESH			
7. PETERED OUT			
8. VERY REFRESHED			
9. FAIRLY WELL POOPED			
10. READY TO DROP			

PREVIOUS EDITION WILL BE USED

SAM FORM SEP 76 136 SUBJECTIVE FATIGUE CHECKCARD

Figure A-1. The USAF School of Aerospace Medicine Subjective Fatigue Checkcard (SAM Form 136). The checkcard is scored by adding 2 points for each check in the "better than" column and 1 point for each check in the "same as" column. Checks in the "worse than" column are not counted.

SLEEP SURVEY

NAME (Last, First, MI)	GRADE	DATE/TIME

1 On the chart below, mark an X in each half hour interval you slept yesterday and today.

YESTERDAY

MID-NIGHT 0100 0200 0300 0400 0500 0600 0700 0800 0900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 MID-NIGHT

TODAY

MID-NIGHT 0100 0200 0300 0400 0500 0600 0700 0800 0900 1000 1100 1200 1300 1400 1500 1600 1700 1800 1900 2000 2100 2200 2300 MID-NIGHT

2. HOW MUCH TROUBLE DID YOU HAVE GOING TO SLEEP LAST NIGHT?

☐ NONE ☐ SLIGHT ☐ MODERATE ☐ CONSIDERABLE

3. HOW RESTED DO YOU FEEL?

☐ WELL RESTED
☐ MODERATELY RESTED ☐ SLIGHTLY RESTED ☐ NOT AT ALL

4. DO YOU FEEL LIKE YOU COULD HAVE USED SOME MORE SLEEP?

☐ YES ☐ NO

SAM FORM SEP 78 154 PREVIOUS EDITION WILL BE USED REMARKS ON REVERSE

Figure A-2. The USAF School of Aerospace Medicine Sleep Survey Form (SAM Form 154).

USAFSAM 48-HOUR MINUTEMAN ALERT TEST
ALERT ACTIVITIES QUESTIONNAIRE

Complete this questionnaire at approximately 1200 while on duty at the LCF. 48-hour crews will complete two for each alert tour.

Name ____________________ Crew # ______ Date ____________ Local Time _________

LCF Designation __________________

Fill in the approximate time you spent in the following activities during the past 24 hours. Round to the nearest half hour. Use zero when little or no time was devoted to an activity.

Traveling to the LCF ______ hrs.

Monitoring Maintenance Activities ________________ hrs.

Maintenance at LCF ___________________________ hrs.

Maintenance at LF ____________________________ hrs.

Processing Security Situations ___________________ hrs.

Processing Message Traffic ___________________ hrs.

Local or Higher HQ
Evaluations/Exercises ____________________________ hrs.

On-site training _____________________________ hrs.

Educational Programs ________________________ hrs.

Other ________ hrs.

Check appropriate response or fill in blanks.

Was your planned sleep schedule disrupted in any way? Yes ___ No ___

If yes, for what reason or reasons. _________________________________

____

Select the term that best describes the workload during the past 24 hrs.

Light __________ Moderate ___________ *Heavy ___________

*Briefly name the activities that affected your choosing this rating.

____

____

Additional remarks on reverse.

Complete the attached fatigue card upon returning to base.

Figure A-3. Alert Activities Questionnaire.

APPENDIX B

ANALYSIS OF RATIOS OF COMPLETION OF SUBJECTIVE FATIGUE CHECKCARDS

Crew members did not complete all response materials during the 90-day test. Table B-1 summarizes the ratios of completed forms to the maximum possible number of Subjective Fatigue Checkcards.

The test group and control group did not differ significantly in their ratios of completion of the pre-alert data. The test group's ratio of completion was significantly greater for the data completed at end-alert, χ^2 (1) = 9.88, $p < .01$; post-return, χ^2 (1) = 4.94, $p < .05$; and next-day, χ^2 (1) = 22.8, $p < .001$. To further test the possibility that missing data were an indication of fatigue, the change in ratios of completion from pre-alert and end-alert were compared in each group. There were significant decreases in the ratios of completion of both the test group, χ^2 (1) = 21.3, $p < .001$, and the control group, χ^2 (1) = 78.4, $p < .001$, for those reporting times. The decrease in the control group's ratios of completion from pre-alert to end-alert was significantly greater than the decrease in the test group's ratios, using an approximate test giving a normal deviate of 2.2, $p < .05$.

TABLE B-1. RATIOS OF SUBJECTIVE FATIGUE CHECKCARDS COMPLETED TO THE MAXIMUM POSSIBLE NUMBER FOR EACH GROUP AND REPORTING TIME

	Group	
Reporting time	Test[a]	Control[b]
Pre-alert	.85	.82
End-alert	.73	.64
Post-return	.54	.47
Next-day	.69	.55

[a]The maximum possible number was 431.

[b]The maximum possible number was 719.

DATE
FILMED
8